DOCUMENT RESUME

ED 199 637 CS 005 921

AUTHOR MacLean, Margaret
TITLE Considering Construct Validity in Incomplete Text Research.
PUB DATE Dec 80
NOTE 18p.; Paper presented at the Annual Meeting of the National Reading Conference (30th, San Diego, CA, December 3-6, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cloze Procedure; Context Clues; *Discourse Analysis; Hypothesis Testing; *Individual Differences; Literature Reviews; Predictor Variables; Reading Achievement; Reading Difficulties; Reading Processes; *Reading Research; *Research Design; *Test Validity

ABSTRACT

A review of research on "incomplete text," which refers to a variety of adaptations of cloze procedures, illustrates how construct validity has been considered and will need to be considered in future research on the subject. Much of the early research on incomplete text focused on product--providing quantitative assessments of comprehension--necessitating efforts to establish reliability and concurrent validity of the measures. As researchers became more involved in manipulating incomplete text, they began increasingly to investigate qualitative aspects of text processing and the issue of construct validity became more important than face validity, concurrent validity, and content validity. The trend has been toward considering text and to a lesser extent reader variables involved in processing information during reading. It appears that incomplete-text researchers need to look beyond simple differences in good and poor readers' cloze scores to determine why differences appear--and to understand how strategies are text or reader related. These insights will be easier to obtain when constructs underlying performance and a sampling of performance (through incomplete-text measures) are identified. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED199637

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Considering Construct Validity in Incomplete Text Research

Margaret MacLean
Reading Centre
McGill University
3700 McTavish
Montreal, Quebec
H3A 1W8

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Margaret MacLean

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS505921

How does and how can incomplete text research provide qualitative insights, within a developmental framework, on the processing strategies of good and poor readers? This paper will discuss these questions by presenting a selective review of research. In particular, the issue of construct validity and how it has been considered by incomplete text investigators will be considered.

The term, "incomplete text," refers to a variety of adaptations of cloze procedures which may or may not conform to traditional notions of cloze. All the research reviewed here uses some variation of incomplete text to examine the product and process of comprehension. However, these variations tend to be quite different, and as a result the familiar "cloze" description suggesting random deletion patterns is not so appropriate.

Much of the early incomplete text research focussed on product - that is, providing quantitative assessments of comprehension. As a result, considerable effort was directed to establishing reliability and concurrent validity of the measures. Using a variety of deletion patterns, scoring procedures, and content materials, a number of researchers, among them, Friedman, 1964; Gallant, 1965; Greene, 1965; Jenkinson, 1957; Ruddell, 1964; Schneyer, 1965, found significantly high correlations between standardized reading tests and incomplete text measures.

Certainly this demonstration of reliability and concurrent validity was an important development in incomplete text research as it provides evidence that insofar as standardized tests measure reading comprehension, so too, incomplete text tasks are also reliable and valid measures.

However, as researchers become more involved in manipulating incomplete text, they have begun increasingly to investigate qualitative aspects of text processing and the issue of construct validity, in contrast to face validity, concurrent validity and content validity become important. How have researchers using incomplete text measures thought of construct validity? How should it be considered? It seems the trend has been toward considering text and to a lesser extent, reader variables involved in processing information during reading. The following sections examine those text and reader variables which have been considered most frequently.

TEXT

Two aspects of text structure which incomplete text researchers have focussed on most often are context and content. That is, researchers have focussed on the influence of context and content variations on incomplete text performance. For the most part, this research has merely provided descriptions of differences in scores without examining how or why these differences appear.

Initially, researchers tended to focus attention on amount and arrangement of context surrounding a deletion. Some early studies indicated that the ability to complete incomplete text is bilaterally constrained by 4 or 5 words in either side of the blank (Aborn, Rubenstein, and Sterling, 1959; MacGinitie, 1961; Miller and Coleman, 1967). However, other researchers have contended that this ability was influenced by context extending beyond the sentence, to the paragraph level, and perhaps even between paragraphs (Horton, 1973; Ramanauskas, 1972; Rankin, 1978).

Attempts to elucidate this issue of the influence of context - whether it was local or global, intra- or intersentence - were made by a series of studies investigating the effect of prior context on incomplete text performance. Suhorsky (1975) varied the amount of context, from 0 - 6 T units, preceding the first deletion in a passage. He found no significant differences in cloze scores of grades 3 and 5 students. Bartoo (1975) examined the effect of variations

of larger amounts of preceding text. He found no significant differences in the performance of grade 10 students on incomplete texts preceded by 0, 300, and 500 words of intact text. Erickson and Hansen (1974) compared the performance of grade 5 students who completed cloze passages embedded in the context of a total message or presented in isolation. They found no significant differences in performances between the in-context and out of context conditions. Further, they found no cumulative effect of context redundancy across passages. Smith-Burke, Gingrich and Eagleeye (1978), investigating the effect of prior context in high school students' performance on incomplete text passages preceded by 0, 250, or 500 words of unmutilated text, found no significant differences in scores due to prior context.

The evidence from these studies appears to be quite convincing. However, one should recall that only prior context was investigated. With the exception of Erickson and Hansen, no one considered the cumulative effect of context redundancy across text. As well, most of the studies discuss results in terms of total scores only. It is possible that some deletions were more sensitive to prior context than others. However, discussion of total scores only, masks any evidence of differential effects of prior context.

Smith-Burke et al., examining individual deleted items, found that certain types of words, especially pronouns and adjectives of quantity, were influenced more by prior context than other words. They suggested that by deleting these or other "cohesive" elements, one could get a more accurate perspective on the effect of contextual buildup.

The notion of selective deletions to assess the influence of context is supported by Rankin and Thomas' (1980) findings. They too found differential effects of context on certain words. In particular, they found that content words were more facilitated by larger context than function words. Their results, in conjunction with Smith-Burke et al., suggest that in order to understand the influence of contextual constraints, one needs to consider more than just total scores and that before concluding that incomplete text measures are not sensitive to larger context, one needs to examine possible differential effects of context and sensitivity of deletions to context.

This point regarding sensitivity of incomplete text measures to context is reiterated by Goetz and Dixon (1979) who replicated Neville and Pugh's (1976-77) investigation of the use of context by good and poor readers on incomplete text tasks where complete or restricted context was available. They found that good readers benefited more from context than poor readers and that poor readers' scores on restricted and complete context passages exhibited much less discrepancy than good readers'. Goetz and Dixon contended that it was difficult to determine from Neville and Pugh's study whether good readers were using extra context from within or across sentences. They examined intra- and intersentence differences in the use of context by good and poor readers by comparing restricted text performance on sentence fragments or whole sentences with complete context performance. They found - as Neville and Pugh did - that good readers made better use of context than poor readers. However, they cautioned this effect was limited to within sentence context only.

It appears that further research, manipulating deletions to maximize dependency on intersentence context is needed. Although previous research has investigated the issue of contextual constraints, it seems that conclusions regarding readers' use or non-use of this larger context have been made without ensuring that deletions are in fact sensitive to constraints operating between sentences or paragraphs. How does one tap "intersentence redundancy?" Should one attempt to identify and delete cohesive items? Cross-text comparisons are needed to provide information on whether contextual constraints operate differently across different styles of texts.

In addition to examining the effect of context on incomplete text performance, researchers have also considered the effect of content, that is, either subject matter or style. Cohen (1975), Miller and French (1974), and McNinch, Kazelskis and Cox (1974) examined the differences in incomplete text scores on passages equated for readability levels, but differing in subject material. Their findings tended to be inconclusive. McKenna (1977), reviewing this literature stated that "...no effect on cloze performance can be definitely ascribed to content area itself..." (p. 40). He does not rule out its possible influence - but merely emphasizes the difficulty in obtaining accurate assessments of the variance due solely to it.

Other researchers have examined content by focussing on style - that is, they have considered differences in incomplete text scores on narrative and expository materials. Smith-Burke et al., who examined

the effect of style on deletions, found significant differences between narrative and expository texts. However, their passages differed in readability levels. The narrative passage, on which students scored significantly higher, had a lower readability level than the expository passage.

Smith-Burke and Gingrich (1979) rectified this problem of differing readability levels in their examination of the effects of style on function and content word replacement. They found no significant difference between narrative and expository texts on function word replacement. However, as the scores were very high, they suspected a ceiling effect was operative. They did find significant differences in content word scores. Content words were more difficult to replace in the narrative than in the expository texts. In discussion of these findings, they explained that the narrative text contained mere synonymy and elaboration than the expository text which was characterized more by repetition and precision of terms and that many of their deletions tapped these repetitions. It could be that their exact word scoring procedure enhanced expository scores over narrative. Perhaps if they had considered synonyms or semantically acceptable replacements, they would have found different patterns of results.

Erickson and Hansen (1974) also examined the effect of style on cloze scores. They found grade 5 students did better on a narrative than an expository passage. They provided no differential discussion

of content/function word replacements, nor any discussion of how the passages may have differed other than style. One might wonder if grade 5 students' familiarity with narrative text materials could have influenced their superior performance in the narrative passage. Also, whether prior knowledge of the narrative topic, "Two Big Bears," in comparison in the expository topic, "The Norwegians of Hemnesberget" could have influenced performance?

The problem with focussing on context and content - of characterizing text in terms of these two aspects - is that we cannot be sure what differences in incomplete text scores mean. Are the differences due to the text or the reader? How different are the narrative and expository texts? How do readers differ in what they know about story grammars in comparison to grammars for exposition? Would researchers find similar differences if they considered a range of narrative texts or a range of expository texts instead of comparing one sample of each?

Is it possible - or desirable - to identify elements common to samples of narrative and expository texts and then examine differences in scores when deletions are based on these elements? Is text structure an important source of variance in examining text processing? What are the different structural and prior knowledge requirements of text? How else might text structure be characterized other than in terms of context and content? It appears that the issue of text structure is one which needs to be described more clearly in order that one can assess its influence on processing.

Reader Variables

Apart from reading ability, incomplete text researchers have tended not to consider reader variables very frequently. Rousch (1972) did examine the relationship between incomplete text scores and background knowledge in a sample of grade 4 readers. He found no significant relationship between the two, although he did suggest that prior knowledge was more useful for processing syntactic than semantic aspects of text. However, one needs to consider these findings carefully as he provided no evidence of reliability of his prior knowledge measure. Also, the fact that his incomplete text measures came from different parts of a long narrative selection and differed in their content/function word ratios may have influenced results.

Rankin and Thomas (1980) examined the influence of passage familiarity on incomplete text performance of grade 8 students. Subjects completed passages either after one reading of the unmutilated text or with no prior exposure to the text. They found that familiarity improved content and function word performance and that its effect was especially significant for content word prediction.

In view of the contradictory findings of these two studies, and the lack of consideration generally in other studies, it would appear that incomplete text researchers should consider how reader variables may influence performance. What reader variables are important to consider? How do reader variables interact with text structure

variables? How can incomplete text measures assess this interaction? Some of the recent research on text processing provides insights.

Processing: Interaction Text and Reader

As early as 1957, Jenkinson considered text and reader variables in her investigation of the product and process of comprehension. Using a standard cloze procedure, she had high school students "think aloud" their responses. With this introspective technique she was able to describe how text and reader variables facilitate text comprehension.

Recently, researchers have begun to modify standard cloze procedures to examine processing strategies. In particular, Cambourne (1977) and DiVesta, Hayward and Orlando (1979) have manipulated traditional cloze to monitor the ongoing silent reading process.

Cambourne deleted words which required the reader to refer backwards or forwards to find clues to meaning; to use the cumulative or logical buildup of text; and to apply real world knowledge which he defined as the network of meanings and relationships known about the topic being read. He found that elementary grade good readers exhibited different behaviours from poor readers. They were able to monitor their reading more effectively, as they referred backwards/forwards more efficiently. They controlled meaning at the "macro-level" better, scoring higher on cumulative text buildup and real

world knowledge deletions than the poor readers. While Cambourne's findings are exciting in that they represent an attempt to manipulate deletions to assess the interaction of text and reader variables, they need to be interpreted cautiously. As he admitted, the deletion categories are rather "fuzzy." That is, one cannot be sure they are tapping what they are supposed to. As well, other deletion categories need to be considered, in case other language processes in addition to/or in place of those sampled in this study are operative during reading.

DiVesta, Hayward and Orlando investigated whether good and poor readers in grades 6, 7, and 8 differed in their use of running or subsequent text. To assess this they designed short passages where information critical to completion of the blanks appeared either before or after the blanks (5 blanks per passage). They found that good and poor readers differed in their use of text. Poor readers were less able than good readers to monitor their comprehension using subsequent text. DiVesta et al. explained that the poor reader may not understand that s/he too must sometimes supply information the text does not, or at least not just rely on serial processing of text information to understand the text. They suggested this change in expectations regarding what the text offers and how to manage it, leading to the monitoring of comprehension "may be an important transition in reading" (p. 105).

While one might critize DiVesta et al.'s reliance on short texts and their use of separate passages for subsequent and running text, their study is an important one as it represents the first attempt to use incomplete text to examine developmental differences in monitoring text for comprehension. As well, along with Cambourne's study, theirs is one of the first to demonstrate how incomplete text measures could be developed to investigate how metacognitive skills operate during text processing.

As incomplete text researchers focus now on processing, more attention needs to be directed towards the issue of construct validity. A number of useful models of text processing and text comprehension exist for building construct validity of incomplete text measures. However, one needs to consider what influence different models might have. What implications for manipulating text deletions arise if one accepts the top-down model of processing? The interactive model? The developmental model? Very few studies have considered developmental aspects when assessing incomplete text performance. Are they important? Consideration and identification of theoretical constructs would facilitate qualitative interpretation of findings. What is good quality performance? How does it differ across ages? Across texts? What influence do text and reader variables have on qualitative insights?

Incomplete text researchers have examined differences in good and poor readers' scores. However, it appears now that one needs to look beyond scores to determine why differences appear - and to understand how strategies are text or reader related. These insights will be easier to obtain when constructs underlying performance and sampling of performance (through incomplete text measures) are identified.

Aborn, M., Rubenstein, H. and Sterling, T. Sources of contextual constraint on words in sentences. Journal of Experimental Psychology, 1959, 57, 171-180.

Bartoo, R. The effect of cumulative context upon cloze task performance of tenth graders. Unpublished doctoral dissertation. State University of New York at Buffalo, 1975.

Cambourne, B. Some psycholinguistic dimensions of the silent reading process. A pilot study. Paper presented at the Annual Meeting of the Australian Reading Conference. Melbourne, Australia, 1977.

Cohen, J. The effect of content area material on cloze test performance. Journal of Reading, 1975, 19, 247-250.

DiVesta, F.J., Hayward, K.G. and Orlando, V.P. Developmental trends in monitoring text for comprehension. Child Development, 1979, 50 (1), 97-105.

Erickson, L. and Hansen, L. Performance in cloze passages in and out of context. In P. Nacke (ed.), Interaction: Research and Practice in College-Adult Reading. 23rd Yearbook NRC. Clemson, South Carolina: NRC, 1974, 158-162.

Friedman, M. The use of the cloze procedure for improving the reading of foreign students at the University of Florida. Unpublished doctoral dissertation, University of Florida, 1964.

Gallant, R. Use of cloze tests as a measure of readability in the primary grades. In J. Allen Figurel (ed.), Reading and Inquiry. IRA Conference Proceedings, 1965, 10, 286-287.

Goetz, E.T. and Dixon, K.M. The use of context by good and poor readers in a cloze task. Paper presented at National Reading Conference, San Antonio, 1979.

Greene, F.P. Modifications of the cloze procedure and changes in reading test performances. Journal of Educational Measurement, 1965, 2 (December), 213-17.

Horton, R.J. The construct validity of cloze procedure: an exploratory factor analysis of cloze, paragraph reading, and structure of intellect. Unpublished doctoral dissertation, Hofstra University, 1973.

Jenkinson, M. Selected processes and difficulties in reading comprehension. Unpublished doctoral dissertation. University of Chicago, 1957.

MacGinitie, W. Contextual constraint in English prose paragraphs. Journal of Psychology, 1961, 51, 121-130.

McKenna, M. An investigation of the cloze procedure as a memory search process. Unpublished doctoral dissertation, University of Missouri - Columbia, 1977.

McNinch, G., Kazelskis, R. and Cox, J.A. Appropriate cloze deletion schemes for determining suitability of college textbooks. In P. Nacke (ed.), Interaction: Research and Practise for College-Adult Reading. 23rd Yearbook of the National Reading Conference. Clemson, S.C.: NRC, 1974.

Miller, G.R. and Coleman, E.B. A set of 36 prose passages calibrated for complexity. Journal of Verbal Learning and Verbal Behaviour, 1967, 6 (December), 851-854.

Miller, W. and French, S. Using the cloze procedure to determine the suitability of social science and science textbooks. In P. Nacke (ed.), Interaction: Research and Practise for College-Adult Reading. 23rd Yearbook of National Reading Conference. Clemson, S.C.: NRC, 1974.

Neville, M.H. and Pugh, A.K. Context in reading and listening: Variations in approach to cloze tasks. Reading Research Quarterly, 1976-77, 12 (1), 12-31.

Ramanauskas, S. The responsiveness of cloze readability measures to linguistic variables operating over segments of text longer than a sentence. Reading Research Quarterly, 1972, 8 (1), 72-91.

Rankin, E. Characteristics of the cloze procedure as a research tool in the study of language. In P. David Pearson and June Hansen (eds.), Reading: Disciplined Inquiry in Reading Process and Practise. 27th Yearbook NRC, Clemson, S.C.: NRC, 1978.

Rankin, E. and Thomas, S. Contextual constraints and the construct validity of the cloze procedure. In M. Kamil and A. Moe (eds.), Perspectives in Reading Research and Instruction. 29th NRC Yearbook. Washington, 1980, 47-55.

Rousch, P.D. A psycholinguistic investigation into the relationship between prior conceptual knowledge, oral reading miscues, silent reading, and post-reading performance. Unpublished doctoral dissertation, Wayne State, 1972.

Ruddell, R. A study of the cloze comprehension technique in relation to structurally controlled reading material. International Reading Association Conference Proceedings, 1964, 9, 298-303.

Schneyer, J.W. Use of the cloze procedure for improving reading comprehension. Reading Teacher, 1965, 19 (December), 174-179.

Smith-Burke, M. and Gingrich, P. Differential role of function words and lexical items in narrative and expository text in M. Kamil and A. Moe (eds.), Reading Research: Studies and Applications. 28th Yearbook of National Reading Conference. Clemson, S.C.: NRC, 1979.

Smith-Burke, M., Gingrich, P. and Eagleeye, D. Differential effects of prior context style, and deletion patterns on cloze comprehension. In P. David Pearson and J. Hansen (eds.). Reading: Disciplined Inquiry in Reading Process and Practice. 27th Yearbook of National Reading Conference. Clemson, S.C.: 1978.

Suhorsky, J. An investigation of the relationship between undeleted text preceding a cloze test and cloze test result. Unpublished doctoral dissertation. University of Maryland, 1975.